Energy Demand of the Road Transport Sector of Saudi Arabia—Application of a Causality-Based Machine Learning Model to Ensure Sustainable Environment

Muhammad Muhitur Rahman 1,* , Syed Masiur Rahman 2,*, Md Shafiullah 3,* , Md Arif Hasan 4, Uneb Gazder 5, Abdullah Al Mamun 6, Umer Mansoor 7, Mohammad Tamim Kashifi 2, Omer Reshi 2, Md Arifuzzaman 1, Md Kamrul Islam 1 and Fahad S. Al-Ismail 2,3,8

1 Department of Civil and Environmental Engineering, College of Engineering, King Faisal University, Al-Ahsa 31982, Saudi Arabia
2 Applied Research Center for Environment & Marine Studies, King Fahd University of Petroleum & Minerals (KFUPM), Dhahran 31260, Saudi Arabia
3 Interdisciplinary Research Center for Renewable Energy and Power Systems, King Fahd University of Petroleum & Minerals, Dhahran 31261, Saudi Arabia
4 Climate Change Response Unit, Wellington City Council, 113 The Terrace, Wellington 6011, New Zealand
5 Department of Civil Engineering, College of Engineering, University of Bahrain, Isa Town P.O. Box 32038, Bahrain
6 Department of Civil and Environmental Engineering, University of Utah, 110 Central Campus Drive, MCE-1435, Salt Lake City, UT 84112, USA
7 Department of Civil and Environmental Engineering, King Fahd University of Petroleum & Minerals (KFUPM), Dhahran 31260, Saudi Arabia
8 Electrical Engineering Department, King Fahd University of Petroleum & Minerals, Dhahran 31261, Saudi Arabia

* Correspondence: mrahman@kfu.edu.sa (M.M.R.); shafiullah@kfupm.edu.sa (M.S.)

Abstract: The road transportation sector in Saudi Arabia has been observing a surging growth of demand trends for the last couple of decades. The main objective of this article is to extract insightful information for the country’s policymakers through a comprehensive investigation of the rising energy trends. In the first phase, it employs econometric analysis to provide the causal relationship between the energy demand of the road transportation sector and different socio-economic elements, including the gross domestic product (GDP), number of registered vehicles, total population, the population in the urban agglomeration, and fuel price. Then, it estimates future energy demand for the sector using two machine-learning models, i.e., artificial neural network (ANN) and support vector regression (SVR). The core features of the future demand model include: (i) removal of the linear trend, (ii) input data projection using a double exponential smoothing technique, and (iii) energy demand prediction using the machine learning models. The findings of the study show that the GDP and urban population have a significant causal relationship with energy demand in the road transportation sector in both the short and long run. The greenhouse gas emissions from the road transportation in Saudi Arabia are directly proportional to energy consumption because the demand is solely met by fossil fuels. Therefore, appropriate policy measures should be taken to reduce energy intensity without compromising the country’s development. In addition, the SVR model outperformed the ANN model in predicting the future energy demand of the sector based on the achieved performance indices. For instance, the correlation coefficients of the SVR and the ANN models were 0.8932 and 0.9925, respectively, for the test datasets. The results show that the SVR is better for predicting energy consumption than the ANN. It is expected that the findings of the study will assist the decision-makers of the country in achieving environmental sustainability goals by initiating appropriate policies.

Keywords: artificial neural network; causality analysis; energy demand; greenhouse gas emission; sustainable environment; machine learning; road transport; Saudi Arabia; support vector regression
1. Introduction

The unprecedented increase in global energy demand requires advanced and comprehensive analyses of energy drivers at national, regional, and international levels. The correlation between energy demand and its independent factors can be understood using econometric methods [1–3]. The results of causal analysis help understand the most important factor of energy demand and assist in accurately estimating future energy demand. It is essential to estimate future energy demand as it has significant policy implications regarding energy security and future economic growth patterns. Recent developments in computing technology, intelligent forecasting methodologies, and algorithms have paved the way for a major breakthrough in modeling and simulation. Machine learning models regress energy demand using socio-economic, demographic, and climatic variables. Typically, these variables are nonlinear. Therefore, energy modeling became a critical issue for practitioners and scientists to contribute to creating sound plans and policies. One of the crucial steps in energy planning is to assess current energy use and forecast future needs [4,5].

Road transportation is one of the major energy consumption sectors in Saudi Arabia [6]. Since renewable energy resources are not used on a large scale in Saudi Arabia for road transportation, it results in continuous increases in the domestic consumption of petroleum products and the associated increase in greenhouse gas (GHG) emissions. The residents of rich urban areas tend to rely on personal automobiles for most of their travel needs [7]. Saudi Arabia has the highest vehicle ownership in the Middle East [8]. The recent royal decree on allowing women to drive will put 9 million potential new drivers on the road [9]. This increasing number of cars will significantly increase the energy demand in road transportation. Future energy demand can be met by analyzing the factors that have historically influenced energy use and making more accurate predictions based on these factors and historical trends. Econometrics techniques can help to find the key drivers from a set of drivers [10]. However, concerning future demand estimation, the traditional regression method cannot address the nonlinearity of different energy demand components. Previous studies suggested adjusting nonlinearity while dealing with energy consumption [11]. Therefore, it is crucial for the Kingdom to investigate the energy demands of road transportation with cutting-edge methods that will enlighten policymakers and decision-makers.

In the relevant literature, the researchers considered different combinations of many variables, which include GDP, gross national product (GNP), population, transport amount (vehicle-kilometer), freight transport amount (ton-kilometer), passenger transport amount (person-kilometer), number of registered vehicles, urbanization rate, and fuel price for developing transport energy models [12–17]. Multivariate, cointegration, and regression analysis can explain the influence of different indicators on energy demand. In Ref. [18], a few selected countries investigated the causal interrelationship between energy consumption and GDP. Ozturk and Acaravchi [19] studied the impact of energy use on GDP growth. Canyurt et al. [20] proposed an energy model using a genetic algorithm, selecting GDP, population, and import and export as inputs. Geem [14] proposed a neural network energy model for South Korea. Azadeh et al. [21] reported a fuzzy regression algorithm based Iranian energy model. Denoised electricity demand data allowed An et al. [22] to isolate the seasonal component and use it to train an ANN model. Uzlu et al. [23] used data on Turkey’s GDP, population, imports, and exports to inform a neural network approach to optimizing the country’s energy demand. Deshani et al. [24] took the first difference of the input series. They used a k-means clustering technique to select ANN model inputs for the prediction of the electricity demand. Table A1 in the Appendix A section summarizes various methodologies employed in literature to model energy demand in general and transportation energy demand. Figure 1 presents the article distribution of different countries worldwide based on the results received in Scopus with the keyword-transportation energy demand from the year 2000 to 2022. In contrast, Figure 2 presents the results obtained with the keywords transportation energy demand and greenhouse gas emissions for
the same period. It was also evident from the search that the number of publications over the years are growing almost exponentially due to the importance of the topic.

Figure 1. Worldwide publication trends obtained from Scopus with the keyword transportation energy demand.

Figure 2. Worldwide publication trends obtained from Scopus with the keywords transportation energy demand and greenhouse gas emissions.

However, according to information presented in Table A1, it was revealed that a significant trend toward applying artificial intelligence (AI) and other regression techniques in energy demand modeling in the transport sector globally. The artificial neural networks approaches are most popular amongst various AI methods. On the other hand, support vector regression, another dominant AI model, has not been explored frequently to model transportation energy demand. However, the SVR requires fewer parameters than other AI models and reaches the global optimum solution at a lower expense. It also does not suffer from the overfitting problem [25]. Due to its benefits, it has been employed in research and industries for decades [26–29]. This approach was successfully used in a wide range of applications such as building energy consumption [30], energy performance [31], solar radiation [32], and electricity load [33]. Successful application of SVR was also found in transportation engineering and planning; for instance, it was used in intelligent transportation systems [34], retro-reflectivity degradation of traffic signs [35], short-term
travel time [36], electric vehicle charging duration time [37], freeway speed [38], and real-
time crash risk on urban expressways [39]. In Saudi Arabia’s context, there is a dearth of
research on energy demand modeling in the transportation sector of Saudi Arabia that
makes use of AI methods. In addition, to the best of the author’s knowledge, the support
vector regression is yet unexplored for Saudi Arabia’s transportation energy modeling.
Considering the mentioned notes, this article develops a causality-based machine-learning
scheme for modeling the energy demand in the transport sector of Saudi Arabia. The
significant contributions of the article are as follows:
• This research employs a vector error correction model (VECM) for causality test
  analysis to identify a short and long-term relationship between the dependent and
  explanatory variables;
• This paper develops an SVR model for forecasting energy consumption in the King-
dom’s road transportation between 2018 and 2030 using a double exponential smooth-
ing method using the projected input dataset.
• Finally, it compares the results from the SVR model with an ANN model to identify a
  suitable model for Saudi Arabia’s energy demand in the transportation sector. The
  ANN model is chosen for comparison purposes due to its popularity in transportation
  sector energy demand modeling.

The following sections of the article are organized as follows: Section 2 provides a
methodology that includes model development data, causality analysis approach, and
prediction approach. Section 3 illustrates the results associated with causality analysis and
energy demand projection models. Finally, the article is concluded in Section 4 with policy
implication-related remarks.

2. Methodology

This section briefly discusses the data used for the development of AI-based energy
demand models in the transport sector of Saudi Arabia. In addition, it introduces the
approaches for causality analysis and energy demand projection. In addition, the step-by-
step procedure for the model development is shown using a flow chart in Section 2.4.

2.1. Model Data

The initial and most critical step for causality test analysis and model development is
the selection of input variables. The data type of the explanatory variables is crucial for
causality test analysis because the VECM requires panel data with trends to identify the
short and long-run relationship between the dependent and explanatory parameters [10].
This study adopted VECM for causality analysis because it is a widely used model for
estimating short- and long-run causal relationships among dependent and explanatory
variables [3,40]. Another reason for using VECM is that it is suitable if some of the ex-
planatory variables are co-integrated among themselves. For example, this study identified
that GDP, fuel price, urban population, and passenger vehicle number are the factors of
transport emissions. Here, GDP and fuel price are co-integrated among themselves as GDP
often depends on fuel price. Given that an explanatory variable is co-integrated among
themselves, VECM can effectively be used for causality tests. Another advantage of using
VECM is that it cannot only estimate the causal relationships between dependent and
independent variables, but it can also estimate causal relationships among the de-pendent
variables themselves [3].

For forecasting energy demand, this study developed some AI-based energy demand
models. We note that the accuracy of the energy demand-forecasting model is highly reliant
on selected input variables. A high number of input variables may lead to overfitting and
overtraining, resulting in reduced accuracy of models. Many socio-economic and demo-
graphic factors like GDP, population size, imports, exports, employment rate, and economic
performance influence energy demand and forecasts. A country’s GDP is correlated with
energy consumption. Some studies reveal that the most important elements of energy
consumption are GDP, imports, exports, population size, fuel price, and the number of
populations in urban agglomeration. Considered an indicator of a nation’s economic health, GDP measures overall economic activity. When GDP rises, people’s living standards rise along with them, leading to increased energy consumption. The amount of energy used is directly related to the number of people since more energy is used as the population grows. Hence, GDP, the number of registered vehicles, total population, fuel price, and the number of populations in urban agglomeration are input variables in this study. Figure 3 shows the trends of some of the variables for Saudi Arabia.

![Figure 3](image-url)

**Figure 3.** The trends of input and output data along with road sector fuel consumption for the period between 1976 and 2017 (Adapted from [41]).

### 2.2. Causality Analysis Approach

Causality analysis based on the VECM has three major steps: unit root test, cointegration test, and short- and long-run Granger causality test. The current study aims to identify the critical drivers of road transport energy consumption, where energy consumption is the dependent variable, and socio-economic variables such as GDP, urban population, fuel price, and vehicle numbers are the explanatory variables. Therefore, the relationship between energy consumption and socio-economic variables can be represented as [3,40]:

\[
E_t = \alpha + \beta_1 GDP_t + \beta_2 P_t + \beta_3 U_t + \beta_4 V_t + \epsilon_t
\]  

(1)

In the above equation, \(E\), \(GDP\), \(P\), \(U\), and \(V\) are road transport energy consumption, annual GDP, fuel price, urban population, and passenger vehicle number, respectively. The parameter \(\alpha\) and \(t\) is intercept and year, respectively. The \(\beta_1\) to \(\beta_4\) are coefficients for \(GDP\), \(P\), \(U\), and \(V\), respectively, and \(\epsilon_t\) is the constant error term.

#### 2.2.1. Unit Root Test

One of the preconditions for using a VECM is that all variables need to be non-stationary at one level and stationary at their first differences. Therefore, the stationary test is crucial in a VECM, requiring a logarithmic form of the previous equation. The logarithmic form can be written as follows [3,40]:
\[
\ln(E_t) = \ln(\alpha) + \beta_1 \ln(GDP_t) + \beta_2 \ln(P_t) + \beta_3 \ln(U_t) + \beta_4 \ln(V_t) + \ln(\varepsilon_t) \tag{2}
\]

The stationarity of the data set may be checked, and the integration sequence of the explanatory factors can be investigated using the Augmented Dickey–Fuller (ADF) and Phillips–Perron (PP) tests. The ADF test is used because it is considered to be the most widely used unit root test [42]. However, according to Azlina et al. [43], the ADF test often fails to reject a unit root. Therefore, this study used the PP and ADF tests for a robust result. The PP test is used because it is suitable mainly for small datasets, and this study has used a small length of time-series data from 1996 to 2017.

2.2.2. Co-Integration Test

Another precondition for developing a VECM is that at least two of the variables used in the study need to be co-integrated. Johansen’s co-integration test investigates the existence of co-integration among variables [44]. This test is used in this study because it is evident from the unit root test that all the variables are integrated in the same order, and the test performs better in such situations [45]. Thus, the Johansen co-integration test’s trace statistic and maximum Eigenvalue statistics are computed to examine the co-integration of the parameters. Trace and maximum Eigenvalue statistics are the two widely used statistics for identifying the number of co-integrating equations [10].

2.2.3. Granger Causality Test

Once it is evident that there is at least one co-integrating equation, a VECM can be developed, and the Granger causality test can be performed using that VECM framework [26]. The Granger causality test examines both the short- and long-term causality and provides the direction of causality. The error correction term (ECT) of a VECM indicates the adjustment speed towards attaining long-run equilibrium and helps to understand the long-run relationship and causality direction between the dependent and explanatory variables. The Wald F statistics of the Granger Causality test are for understanding the causal relationships and their directions among variables in the short run because this is the most widely used technique to explain short-run relationships [10].

2.3. Prediction Approach

The five main steps of the adopted methodology for future energy demand estimation include (i) input analysis, (ii) input projection, (iii) model construction, (iv) model testing, and (v) output (road transportation energy consumption in Saudi Arabia) forecasting:

Step 1: Input Processing: All the input data exhibit an increasing trend. Each variable’s linear trend is removed by developing a first-order linear regression model with the year as an input. A set of modified variables is obtained by eliminating the linear trend. In the next stage, normalization is performed to scale the data within the range between \(-1\) and \(1\);

Step 2: Input Projection: A double exponential smoothing technique is employed to project the input data between 2018 and 2030;

Step 3: Model Development: The model is developed using the normalized input data between 1976 and 2004;

Step 4: Model Testing: The model is tested for the training data between 1976 and 2004 and the testing data between 2004 and 2017, considering suitable error measures;

Step 5: Output Forecasting: The model is utilized to make projections regarding the output. (road transport energy consumption in Saudi Arabia) for the period between 2018 and 2030 using the projected input data.

2.3.1. Double Exponential Smoothing

The analysis of the GDP, the number of registered vehicles, the total population, and the population in urban agglomeration data reveals a long-run trend. Double exponential smoothing is preferred over single exponential smoothing because it takes into account
both the average and pattern. For n-periods-ahead prediction ($F_{t+n}$), the double exponential forecasting equation is as follows [46–48]:

$$F_{t+n} = P_t + nb_t$$  \hspace{1cm} (3)

where $P_t$ is the projected intercept; $b_t$ is the projected slope. The equations are as follows [46–48]:

$$P_t = ax_t+(1-a)(P_{t-1}+b_{t-1}) \text{ for } 0 \leq a \leq 1$$  \hspace{1cm} (4)

$$b_t = \gamma (P_t - P_{t-1}) + (1 - \gamma) b_{t-1} \text{ for } 0 \leq \gamma \leq 1$$  \hspace{1cm} (5)

where $[x_t]$ represents the raw data sequence, $a$ and $\gamma$ are the data and trend smoothing factors, respectively. The selected initial values of $P_1$ and $b_1$ are equal to the observed value at $t = 1$ (i.e., $x_1$) and the difference between the observed values at $t = 2$ and $t = 1$ (i.e., $x_2 - x_1$), respectively.

### 2.3.2. Support Vector Regression

The SVR was initially proposed by Cortes and Vapnik [49] on the basis of the structural risk minimization principle. The technique allows us to reduce the generalization error constraint without worrying too much about the training error. Generalization error is critical in evaluating the algorithm’s accuracy in forecasting unseen data. Although the applications of SVR were initially restricted to pattern recognition problems, the regression problems can also be solved now. The SVR builds an optimal geometric hyperplane for separating the data. It also uses nonlinear mapping ($\phi$) to transform the data into a high-dimensional feature set before performing the linear regression in the transformed feature space [50–52].

For a mathematical explanation of the concept, let us consider $x \in R^n$ and $y \in R$, the hyper-plane function, $y = f(x) = w.\phi(x) + b$, where $w \in R^n$ = weight vector, and $c \in R$ = bias. The function $\phi(x)$ is a nonlinear transformation from $R^n$ to a higher dimensional space. Now, it is required to discover the $w$ and $b$ values for the determination of the $x$ values by minimizing the regression risk [49–52]:

$$R = \frac{1}{2} \sum_{i=1}^{n} (f(x_i) - y_i)^2 + \frac{\lambda}{2} \parallel w \parallel^2$$  \hspace{1cm} (6)

where $\lambda$ is the regularization constant, $n$ indicates the sample inputs $(x_1, . . . . . . , x_n)$, $(y_1, . . . . . . , y_n)$ refers to target output, and $w$ represents the optimal desired weights vector of the regression hyperplane and can be represented as:

$$w^* = \sum_{i=1}^{N} (\beta_i - \beta_i^*) \phi(x_i)$$  \hspace{1cm} (7)

where $\beta_i, \beta_i^*$ are the solutions to the mentioned quadratic equation [51]. The regression equation can be rewritten by substituting $w^*$:

$$f(x, \beta, \beta^*) = \sum_{i=1}^{N} (\beta_i - \beta_i^*) (\phi(x_i) \phi(x_j) + b) = \sum_{i=1}^{N} (\beta_i - \beta_i^*) K(x_i, x_j) + b$$  \hspace{1cm} (8)

Here, $K(x_i, x_j)$ is known as the kernel function, which is the product of the vectors $x_i$ and $x_j$ in the feature space. The considered kernel function can be a linear, polynomial, or radial basis function (RBF) and is written as:

$$K(x_i, x_j) = \exp \left( -\frac{\parallel x_i - x_j \parallel^2}{2\sigma^2} \right)$$  \hspace{1cm} (9)

where the user provides $\sigma$.

One common form of a cost function is Vapnik’s $\varepsilon$-insensitive loss function [53]:
\[ E_{\epsilon}(f(x_i) - y_i) = \begin{cases} |f(x_i) - y_i| - \epsilon, & \text{for } |f(x_i) - y_i| \geq \epsilon \\ 0 & \text{Otherwise} \end{cases} \]  

(10)

Now, the quadratic problem is defined as:

\[
\text{Minimize } \frac{1}{2} \sum_{i=1}^{n} \sum_{j=1}^{n} (\beta_i - \beta_j^*) (\beta_j - \beta_j^*) K(x_i, x_j) - \varepsilon \sum_{i=1}^{n} (\beta_i + \beta_i^*) + \sum_{i=1}^{n} (\beta_i - \beta_j^*) y_i \\
\text{Subject to: } \sum_{i=1}^{n} (\beta_i - \beta_j^*) = 0, 0 \leq a_n \leq C, 0 \leq a_n^- \leq C 
\]  

(11)

The \( \beta_i \) and \( \beta_j^* \) are the forces which push and pull the estimate \( f(x_i) \) towards the target output \( y_i \) [54]. The constant \( C \) values cause penalties for errors in estimation for balancing the training error and generalization capability. The biases are determined using the Karush–Kuhn–Tucker (KKT) conditions [49–52].

2.3.3. Artificial Neural Networks

An ANN is a system of interconnected computing nodes (neurons). The information is processed similarly to how the human brain would. Multilayer perceptron (MLP) networks are the most popular type of neural networks. Each layer consists of neurons. Every neuron has a different weight associated with it. The information is passed from the input layer through hidden layers and finally reaches the output layer. Every neuron except the neuron in the input layer receives the information from the neurons in the preceding layer. After this, the neuron passes information to the output through a sigmoid function [55–58]. A training algorithm is adopted to obtain the weights while the algorithm minimizes the cost function, such as mean squared error considering the target and the model output. A general representation of a neural network is shown in Figure 4. As can be seen, the network consists of two inputs, one output, and one hidden layer. The input layer is not associated with any calculations; it simply transfers the input to the first hidden layer, while the remaining connections carry real-valued connection weights that modify the signal strength carried by other nodes. As inputs, the node of hidden layers and output layer receives the sum of the previous layer’s weighted outputs and the bias. The corresponding activation function modifies the input and transfers the result to the nodes of the subsequent layer or the environment.

\begin{figure}[h]
\centering
\includegraphics[width=0.5\textwidth]{figure4.png}
\caption{A simplified illustration of an ANN.}
\end{figure}

The output of a node in the hidden layer and the output layer are in the following part as adopted from [55–58]. The output of the hidden node \( j \):

\[
z_j = f(q) = \frac{e^q}{e^q + 1}, \text{ where } q = \sum_i w_{ij} x_i + w_{oij} 
\]  

(13)
where $w_{ij}$ is the weight of the connection from the $i$th input node to the $j$th hidden node, and $w_{0j}$ is the bias of the $j$th hidden node. In the preceding example, the activation function of the hidden node is a sigmoid function with an input of $q$.

The final output can be represented as [56,57]:

$$y = w_{35}z_1 + w_{45}z_2$$

(14)

where $w_{35}$ and $w_{45}$ are the weights, and $z_1$ and $z_2$ are the output of the third and fourth nodes, respectively.

2.4. Model Development

The ANN and SVR models were developed using MATLAB software. Data sets from 1976 to 2017 were used for analysis. Data from 1976–2004 were utilized for training, whereas data from 2005–2017 were used for testing. The data were projected up to 2030 using a double exponential smoothing technique. Data and trend smoothing factors were selected using the systematic trial and error process. Road transportation energy consumption was modeled using the projected input. Two important hyperparameters which need to be specified in the SVR model are appropriate $C$ and $\epsilon$ values. With the goal of maintaining a balance between the learning error and the complexity of the model, $C$ determines the optimal number of support vectors. The lower and higher $C$ values are associated with underfitting and overfitting, respectively. When used to SVR, the $\epsilon$ is a normalization setting that establishes a compromise between error margin and model robustness to produce optimal adaptation on a new testing dataset. The smoothness of SVR’s response is affected by the value of $\epsilon$. As a result, the model’s complexity and predictive validity rely on the parameter value of $\epsilon$. After many systematic trial and error experiments, the values of $C$ and $\epsilon$ are fixed to 1000 and 0.001, respectively. The conceptual framework of the suggested AI-model for predicting Saudi Arabia’s energy needs for road transport is depicted in Figure 5.

![Conceptual illustration of the proposed AI-model for Saudi Arabia’s energy demand forecasting in the road transportation sector.](image)

Figure 5. Conceptual illustration of the proposed AI-model for Saudi Arabia’s energy demand forecasting in the road transportation sector.

The important hyperparameters of an ANN include the number of hidden layers, the number of neurons in each layer, the learning algorithm, the activation function, the
learning rate, and the learning goal. Due to the small dataset, only one hidden layer was considered in this study. The number of neurons in the hidden layer was ascertained through a systematic trial and error approach considering 1 to 4 neurons. The mean absolute percentage error values for the training dataset varied between 4.34% and 9.35% for the considered neurons. The testing results varied between 4.82% and 5.26% for the same number of neurons. The ANN with three neurons in the hidden layer with an activation function of tan-sigmoid, a learning rate of 0.00002, and a goal of 0.000001 produce the best result considering both the training and the testing results. Figure 6 presents the comparative analysis of ANN and SVR models in terms of mean squared error (MSE) index where the minimum values for the models were 0.0295 (MTOE) and 0.0001 (MTOE), respectively.

Figure 6. Comparative illustrations of the ANN and SVR models in terms of MSE index.

3. Results and Discussion

The section starts by demonstrating the results obtained from causality analysis approaches. Then, it presents the result of the developed SVR-based energy demand model along with different statistical performance measures to verify the efficacy of the developed model. Finally, a comparative analysis between the SVR-based and ANN-based approaches is also discussed.

3.1. Causality Analysis Model Results

In the unit root test, two specifications are used: the ‘intercept’ and the ‘Intercept and Trend.’ Table 1 presents the ADF and the PP test results for both specifications. All the variables are non-stationary at the level and stationary at their first differences, which is the first condition for developing a VECM for causality test analysis. The second condition for creating a VECM is the existence of cointegration between at least two of the variables. The Johansen cointegration test is performed to understand the cointegration among variables. The result is presented below in Table 2. The Johansen cointegration test result based on the Trace statistics and the Maximum Eigenvalue statistics shows that at least one co-integrating equation exists at a 1% significance level. The summary of the co-integrating equation is presented in Table 3. The signs of the coefficient values of the explanatory variables indicate that GDP and the number of vehicles have a positive relationship with transport energy consumption. In contrast, the relationship is negative between the urbanization rate and fuel prices. An increase in GDP or vehicle number increases transport energy consumption, while an increase in urbanization rate or fuel price decreases energy consumption. Since cointegration exists among variables and all the variables are non-stationary ‘at level’ and stationary ‘at their first differences’, a VECM has been developed to understand the
long- and short-run causal relationship between transport energy consumption and its explanatory variables, as can be seen in Table 4.

Table 1. The unit root test result.

<table>
<thead>
<tr>
<th>Variable</th>
<th>At Level</th>
<th>ADF (Lag Length)</th>
<th>PP (Lag Length)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Intercept</td>
<td>Intercept and Trend</td>
<td>Intercept</td>
</tr>
<tr>
<td>E</td>
<td>2.03 (1)</td>
<td>−2.23 (0)</td>
<td>2.42 (3)</td>
</tr>
<tr>
<td>GDP</td>
<td>−0.50 (3)</td>
<td>−1.92 (0)</td>
<td>−0.50 (6)</td>
</tr>
<tr>
<td>P</td>
<td>−1.16 (3)</td>
<td>−1.04 (0)</td>
<td>−1.43 (1)</td>
</tr>
<tr>
<td>U</td>
<td>4.14 (3)</td>
<td>−3.66 ** (4)</td>
<td>4.65 (2)</td>
</tr>
<tr>
<td>V</td>
<td>1.36 (1)</td>
<td>−0.68 * (1)</td>
<td>7.09 (2)</td>
</tr>
<tr>
<td></td>
<td>At first difference</td>
<td>Intercept</td>
<td>Intercept and Trend</td>
</tr>
<tr>
<td>E</td>
<td>−4.81 (0) ***</td>
<td>−6.19 (0) ***</td>
<td>−4.82 (2) ***</td>
</tr>
<tr>
<td>GDP</td>
<td>−3.82 (2) ***</td>
<td>−3.75 (0) **</td>
<td>−3.84 (2) ***</td>
</tr>
<tr>
<td>P</td>
<td>−3.43 (0) **</td>
<td>−3.39 (0) *</td>
<td>−3.43 (0) **</td>
</tr>
<tr>
<td>U</td>
<td>−2.50 * (4)</td>
<td>2.34 (3)</td>
<td>−2.07 * (2)</td>
</tr>
<tr>
<td>V</td>
<td>−0.39 (0) *</td>
<td>−2.00 (0)</td>
<td>−0.39 (0)</td>
</tr>
</tbody>
</table>

Note: E, GDP, P, U, and V are road transport energy consumption, annual GDP, fuel price, urban population, and passenger vehicle number, respectively. The numbers in parentheses are lag lengths, while the asterisks (***, **, *) denote statistical significance at the 0.01, 0.05, and 0.1 levels, respectively.

Table 2. Results of the Johansen cointegration test.

<table>
<thead>
<tr>
<th>Hypothesized Number of Co-integrating Equation(s)</th>
<th>r = 0</th>
<th>r = 1</th>
<th>r = 2</th>
<th>r = 3</th>
<th>r = 4</th>
</tr>
</thead>
<tbody>
<tr>
<td>Trace statistics</td>
<td>113.91 ***</td>
<td>61.58 ***</td>
<td>27.04</td>
<td>7.21</td>
<td>0.57</td>
</tr>
<tr>
<td>Maximum Eigenvalue statistics</td>
<td>52.33 ***</td>
<td>34.54 ***</td>
<td>19.83</td>
<td>6.64</td>
<td>0.57</td>
</tr>
</tbody>
</table>

Note: *** indicate significance at 0.01 level.

Table 3. Summarized form of the co-integrating equation.

<table>
<thead>
<tr>
<th>Dependent Variable: TE</th>
<th>Explanatory Variables</th>
<th>Coefficients</th>
<th>Standard Error</th>
<th>t-Statistics</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Constant</td>
<td>11.8</td>
<td>13.71</td>
<td>0.86</td>
</tr>
<tr>
<td></td>
<td>GDP</td>
<td>0.003</td>
<td>0.017</td>
<td>0.176</td>
</tr>
<tr>
<td></td>
<td>P</td>
<td>−127.10</td>
<td>35.91</td>
<td>−3.54</td>
</tr>
<tr>
<td></td>
<td>U</td>
<td>−17.73</td>
<td>2.53</td>
<td>−7.00</td>
</tr>
<tr>
<td></td>
<td>V</td>
<td>14.35</td>
<td>1.82</td>
<td>19.23</td>
</tr>
</tbody>
</table>

The Granger causality test result shows that GDP and urbanization rate have short-run and long-run-causal relationships with transport energy consumption. However, the number of vehicles and fuel prices do not have a noteworthy causal relationship with energy consumption. Although vehicle number does not have a short-run causal relationship with fuel consumption, there is a causal relationship running from vehicle number to urban population rate and GDP. This means that policy measures are required to promote alternative modes of transportation, especially in urban areas. Since GDP has a strong causal relationship with transport energy consumption in the short and long run, energy intensity needs to be immediately reduced through appropriate policy measures. In addition, urbanization needs to be managed through high-density urban development instead of urban sprawl.
### Table 4. Granger causality test result.

<table>
<thead>
<tr>
<th></th>
<th>Short-Run Granger Causality—F Statistics</th>
<th>Long-Run Granger Causality—t-Statistics</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Ln (E)</td>
<td>Ln (GDP)</td>
</tr>
<tr>
<td>Ln (E)</td>
<td>-</td>
<td>3.93 * (0.07)</td>
</tr>
<tr>
<td>Ln (GDP)</td>
<td>1.52 (0.24)</td>
<td>-</td>
</tr>
<tr>
<td>Ln (P)</td>
<td>3.11 (0.10)</td>
<td>0.01 (0.92)</td>
</tr>
<tr>
<td>Ln (U)</td>
<td>0.60 (0.45)</td>
<td>0.29 (0.60)</td>
</tr>
<tr>
<td>Ln (V)</td>
<td>0.003 (0.96)</td>
<td>2.17 (0.17)</td>
</tr>
</tbody>
</table>

Note: p-values are shown in parentheses, and the symbols ** and * denote statistical significance at the 0.05 and 0.1 levels, respectively.

### 3.2. AI-Based Model Results

The input variables are forecasted to project the road transportation energy demand for Saudi Arabia between 2017 and 2030 (Figures 7 and 8). Both the training (1976–2004) and testing (2005–2017) datasets are used to assess the created model’s efficacy. The model appears to produce results near the observed values for both training and testing datasets. Figure 9 presents the comparative illustrations of the machine learning models predicted energy demand for the road transportation sector of Saudi Arabia with actual energy demand. As can be seen, the SVR model predicted numbers are much closer to the actual numbers than ANN predicted numbers. Therefore, the superiority of the SVR model over the ANN model for the case under study (Saudi Arabia’s transportation sector energy prediction) is verified. The effectiveness of the model is assessed with the help of the root mean squared error (RMSE), mean absolute error (MAE), mean absolute percentage error (MAPE), correlation coefficient (CC), and Willmott’s index of agreement (IA). The values of IA for an ideal match and complete disagreement are 1 and 0, respectively [59]. The created model shows satisfactory performance on the training and testing datasets, as shown by the estimated values of the investigated performance measures (Table 5). The comparative analyses of the achieved performance indices for both machine-learning models are presented in Figure 10.

![Figure 7. Energy consumption from road transport and model output of the neural network.](image)
Figure 8. Road transport energy consumption and model output of the SVR.

Figure 9. Comparative illustrations of the actual versus machine learning model predicted energy demand of the road transportation sector of Saudi Arabia.
Table 5. Performance measures of developed ANN and SVR models.

<table>
<thead>
<tr>
<th>Heading</th>
<th>RMSE (MTOE)</th>
<th>MAE (MTOE)</th>
<th>MAPE</th>
<th>CC</th>
<th>IA</th>
</tr>
</thead>
<tbody>
<tr>
<td>Training Datasets (ANN)</td>
<td>0.7</td>
<td>0.43</td>
<td>4.34</td>
<td>0.9924</td>
<td>0.9989</td>
</tr>
<tr>
<td>Testing Datasets (ANN)</td>
<td>3.9</td>
<td>15.39</td>
<td>7.53</td>
<td>0.8932</td>
<td>0.9899</td>
</tr>
<tr>
<td>Training Datasets (SVR)</td>
<td>0.0</td>
<td>0.00</td>
<td>0.00</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Testing Datasets (SVR)</td>
<td>0.9</td>
<td>0.73</td>
<td>1.91</td>
<td>0.9925</td>
<td>0.9996</td>
</tr>
</tbody>
</table>

Figure 10. Comparisons of the performance measures (CC and IA) in predicting future transportation energy demand for the developed machine learning models.

Figure 11 is a scatter plot illustrating the correlation between the data and the model’s predictions. An identity line, i.e., a \( y = x \) line, is generally drawn as the reference. The data points coincide with the identity line whenever the model’s predictions and observations are in perfect numerical agreement. For both the test and training datasets, the scatter diagram of the observable data and the model output showed that the prediction model reasonably conforms to the observed data. The coefficients of determination \( (R^2) \) values are also satisfactory. Removing the trend line to concentrate on the nonlinear aspect of energy consumption and employing the twofold smoothing procedure for given input projection may account for the reliability of the observed data and the model’s predictions.

Figure 11. Cont.
4. Conclusions and Policy Implementation

The rising energy usage of Saudi Arabia’s transport sector poses challenges to policymakers in achieving the country’s sustainability goals. As a result, the Kingdom must identify the factors that influence energy consumption trends in the sector and develop mitigation plans without compromising the country’s development. Considering the mentioned note, this article studied a few selected socio-economic drivers (GDP, fuel price, urban population, and passenger vehicle number) and their impacts on the energy trends of the sector using econometric analysis. Then, it predicted the future energy demand of the sector employing machine learning tools.

The econometric model analysis revealed that the GDP and urban population have short- and long-run causal relationships with road transport energy consumption. As of the co-integration test, the GDP affected energy consumption negatively, while urbanization affected it positively. Therefore, Saudi Arabia needs to shift its GDP growth towards a low-energy-intensive economy and manage its urban population growth tactfully to reduce both short- and long-run road transport energy demand. In addition, road transportation’s greenhouse gas emissions are directly proportionate to energy consumption because fossil fuels are the only source of energy in Saudi Arabia. Investments in renewables, public transport infrastructures, and low-carbon service sectors could be crucial to breaking the nexus between GDP, urbanization rate, and road transport energy consumption. In addition, promoting electric vehicles powered by renewables and fuel-efficient cars can be a valid option to reduce transport fuel consumption. Shift urban sprawl to smart growth to decouple the urban population from road transport energy consumption. Such a shift is likely to promote active travel (walking, cycling, etc.) and support public transport services, thereby reducing road transport energy consumption. These policy options are crucial for Saudi Arabia to tackle the critical drivers of energy demand in the transport sector and achieve environmental sustainability.

Prediction results of the future energy demand of road transportation employing machine learning models validated the efficacy of the SVR model over the ANN model. All selected statistical performance measures (RMSE, MAE, CC, and IA) for both training and testing datasets were better for the SVR model than the ANN model. For instance, the IA for the ANN model was 0.9899 for the testing dataset, whereas the value was only 0.9996 for the SVR model. Therefore, the concerned authorities can utilize the model for scenario development using different policy approaches to curtail road energy consumption. Although the overall performance of both machine learning models was adequate, both models did not realize the recent changes in energy demand well. Including additional relevant variables with longer time-series data can enhance the model performance. Moreover, other promising machine learning techniques including transparent machines, deep learning
models, etc. can also be explored as an extension of this research to model energy demand of various sectors of the country.


**Funding:** This research was funded by the Deanship of Scientific Research at King Faisal University (KFU), Al-Ahsa 31982, Saudi Arabia, through Project No. GRANT1517.

**Institutional Review Board Statement:** Not applicable.

**Informed Consent Statement:** Not applicable.

**Data Availability Statement:** The data that support the findings of this study are available from the corresponding author, M.M.R. (mrahman@kfu.edu.sa), upon reasonable request.

**Acknowledgments:** The authors acknowledge the support received from the Deanship of Scientific Research at King Faisal University (KFU), King Fahd University of Petroleum & Minerals (KFUPM), and Al-Imam Mohammad Ibn Saud Islamic University, Saudi Arabia, University of Bahrain, and University of Utah, USA.

**Conflicts of Interest:** The authors declare no conflict of interest.

**Appendix A**

### Table A1. Review of the energy demand estimation in the road transport sector.

<table>
<thead>
<tr>
<th>Country and Ref.</th>
<th>Methodology</th>
<th>Model Data</th>
<th>Forecasting Period</th>
<th>Predicted Variable</th>
<th>Predictor Variables</th>
</tr>
</thead>
<tbody>
<tr>
<td>Greece by Polemis in 2006 [63]</td>
<td>Log-linear model</td>
<td>1978–2003</td>
<td>Not specified</td>
<td>Diesel demand</td>
<td>Per capita income and vehicle fleet, and gasoline and diesel prices</td>
</tr>
<tr>
<td>South Korea by Geem et al. in 2007 [64]</td>
<td>ANN</td>
<td>1980–2007</td>
<td>Not specified</td>
<td>Energy demand</td>
<td>GDP, import, export, and population</td>
</tr>
<tr>
<td>Turkey by Szejnor et al. in 2007 [66]</td>
<td>ANN</td>
<td>1968–2005</td>
<td>Not specified</td>
<td>Net energy demand</td>
<td>GDP, population, import, and export</td>
</tr>
<tr>
<td>Country and Ref.</td>
<td>Methodology</td>
<td>Model Data</td>
<td>Forecasting Period</td>
<td>Predicted Variable</td>
<td>Predictor Variables</td>
</tr>
<tr>
<td>-----------------</td>
<td>-------------</td>
<td>------------</td>
<td>--------------------</td>
<td>--------------------</td>
<td>---------------------</td>
</tr>
<tr>
<td>Pakistan by Shabbir et al. in 2010 [70]</td>
<td>LEAP</td>
<td>2000</td>
<td>2000–2030</td>
<td>Urban passenger transport energy consumption</td>
<td>Total number of vehicles, VMT, occupancy level, modal split, and fuel efficiency</td>
</tr>
<tr>
<td>Croatia by Puklec et al. in 2013 [73]</td>
<td>Energy demand of transport (EDT) model</td>
<td>Not specified</td>
<td>2008–2050</td>
<td>Long-term energy demand</td>
<td>Railway, road, seawater and coastal, inland waterway, and air transports</td>
</tr>
<tr>
<td>Thailand by Tansawat et al. in 2015 [76]</td>
<td>Linear and log-linear regression models</td>
<td>2007</td>
<td>Not specified</td>
<td>Transport fuel consumption</td>
<td>Gross provincial product, total number of sedans and gas stations, and a few dummy variables</td>
</tr>
<tr>
<td>China by Teng et al. in 2017 [80]</td>
<td>Group method of data handling</td>
<td>1980–2011</td>
<td>2012–2052</td>
<td>Long-term energy demand</td>
<td>GDP, population, urbanization rate, incomes, passenger, and freight turnovers, registered vehicle numbers, and fuel retail price index</td>
</tr>
<tr>
<td>24 countries in the Latin America by Llorca et al. in 2017 [81]</td>
<td>Stochastic frontier approach</td>
<td>1990–2010</td>
<td>Not specified</td>
<td>Energy demand</td>
<td>GDP, population, energy price index, gross value added, and population density</td>
</tr>
<tr>
<td>China by Peng et al. in 2018 [82]</td>
<td>CREG model</td>
<td>2015</td>
<td>2015–2050</td>
<td>Energy demand and GHG emissions</td>
<td>GDP, population, vehicle miles traveled, and vehicle stock</td>
</tr>
<tr>
<td>Indonesia by Deendarlanto et al. in 2020 [84]</td>
<td>Sustainable mobility project (SMP) model</td>
<td>1999 –2013</td>
<td>2014–2030</td>
<td>Energy demand</td>
<td>GDP, population, vehicle type, travel distance, and fuel economy</td>
</tr>
<tr>
<td>Turkey by Talebi et al. in 2021 [85]</td>
<td>ANN</td>
<td>1975–2016</td>
<td>2020–2030</td>
<td>Energy demand</td>
<td>GDP, population, oil prices, ton-km, vehicle-km, and passenger-km</td>
</tr>
<tr>
<td>Taiwan by Yao et al. in 2021 [86]</td>
<td>Convolutional neural network</td>
<td>1999–2019</td>
<td>Up to 2025</td>
<td>Energy demand</td>
<td>GDP, population, number of registered vehicles, passenger transport value, and oil price</td>
</tr>
<tr>
<td>Turkey by Turgut et al. in 2021 [89]</td>
<td>OPTSGULL algorithm</td>
<td>1970–2017</td>
<td>2018–2028</td>
<td>Energy demand</td>
<td>GDP, population, employment, trade, inflation, crude oil price, total amount of goods transported, and total vehicle travel in km</td>
</tr>
<tr>
<td>Turkey by Ozdemir and Dörterler in 2022 [90]</td>
<td>Linear, exponential, and quadratic models assisted by heuristic algorithm</td>
<td>1970–2013</td>
<td>2014–2034</td>
<td>Energy demand</td>
<td>GDP, total vehicle kilometer/year, population</td>
</tr>
</tbody>
</table>
Table A1. Cont.

<table>
<thead>
<tr>
<th>Country and Ref.</th>
<th>Methodology</th>
<th>Model Data</th>
<th>Forecasting Period</th>
<th>Predicted Variable</th>
<th>Predictor Variables</th>
</tr>
</thead>
<tbody>
<tr>
<td>Morocco Oubnaki et al. in 2022 [80]</td>
<td>Regression models</td>
<td>1990–2014</td>
<td>2020–2030</td>
<td>Energy demand</td>
<td>GDP, population, urbanization, fuel price, working women rate, number of vehicles registration and active vehicles on the road, and activity rate by gender and category</td>
</tr>
<tr>
<td>28 European countries by Maaouane et al. in 2022 [91]</td>
<td>ANN</td>
<td>1990–2019</td>
<td>2020–2050</td>
<td>Energy demand</td>
<td>GDP, population density, gasoline and diesel price, purchasing power parity, price index, and household final consumption expenditure</td>
</tr>
<tr>
<td>Turkey by Sahraei and Çodur in 2022 [86]</td>
<td>Hybrid meta-heuristic ANN</td>
<td>1975–2019 (First 80%)</td>
<td>1975–2019 (Last 20%)</td>
<td>Energy demand</td>
<td>GDP, population, oil price, passenger-km, vehicle-km, and ton-km</td>
</tr>
</tbody>
</table>

References


35. Jamal, A.; Reza, I.; Shafiullah, M. Modeling Retroreflectivity Degradation of Traffic Signs Using Artificial Neural Networks. *IATSS Res.* in press. [CrossRef]


52. Shaﬁullah, M.; Abido, M.A.; Al-Mohammed, A.H. Artiﬁcial Intelligence Techniques. In Power System Fault Diagnosis; Elsevier: Cambridge, MA, USA, 2022; Chapter 3; pp. 69–100. [CrossRef]


63. Polemis, M. Empirical Assessment of the Determinants of Road Energy Demand in Greece. Energy Econ. 2006, 28, 385–403. [CrossRef]

64. Geem, Z.; Roper, W. Energy Demand Estimation of South Korea Using Artificial Neural Network. Energy Policy 2009, 37, 4049–4054. [CrossRef]


72. Ahmadian, M.; Dastjerdi, A.; Aminekooei, K. General Procedure for Long-Term Energy-Environmental Planning for Transportation Sector of Developing Countries with Limited Data Based on LEAP (Long-Range Energy Alternative Planning) and EnergyPLAN. Energy 2014, 77, 831–843. [CrossRef]


86. Sahraei, M.A.; Çodur, M.K. Prediction of Transportation Energy Demand by Novel Hybrid Meta-Heuristic ANN. Energy 2022, 249, 123735. [CrossRef]


